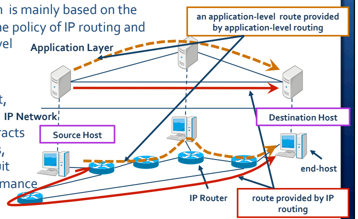
Centralized and distributed heuristic algorithms for application-level traffic routing

○Kazuhito MATSUDA†, Go HASEGAWA†, Satoshi KAMEI†† and Masayuki MURATA†

† Graduate School of Information Science and Technology, Osaka University
†† Service Integration Laboratories, Nippon Telegraph and Telephone Corporation

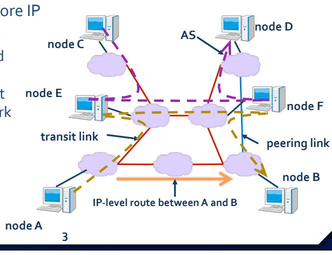
Background: Application-level traffic routing

- Routing mechanism worked on application layer
  - It uses user-perceived routing metrics such as end-to-end latency and available bandwidth
  - It can improve user-perceived performance using detour route traversing another end-host
    - The performance gain is mainly based on the difference between the policy of IP routing and that of application-level traffic routing
    - IP routing is based on router-level hop count, AS-level hop count and commercial contracts with neighboring ISPs, which is not always suit user-perceived performance



Problems in application-level traffic routing

- If a user chooses an application-level (AL) route ignoring other users' AL routes, some AL routes share the same AL link
  - It leads to performance degradation of AL routes, for example increasing end-to-end latency and decreasing available bandwidth
- AL routes may include more IP links than IP-level routes
  - The number of traversed transit links increases, which means that transit cost in the whole network increases

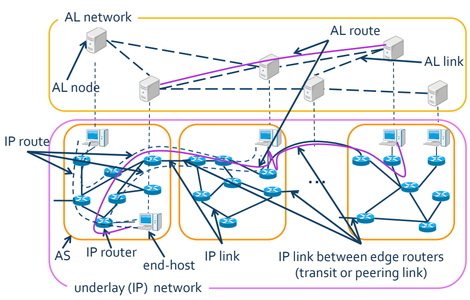


Our goal

Propose an application-level traffic routing operated in a coordinated manner by all AL nodes

1. Formulate the application-level traffic routing as an optimization problem
2. Propose heuristic algorithms to obtain near-optimal solutions to the problem
   - We propose both centralized and distributed algorithms to accommodate wide-range application scenario

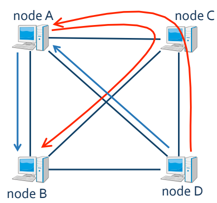
Network model



AL routes optimization problem

- Find the set of AL routes, each of which provides optimal performance for the AL node pair

  Ex) Find the set of AL routes that provide the minimum latency on average
  - AL route must be chosen considering usages of AL links by any other AL routes
    - Because the AL link usages by AL routes affect performance of routes each other
  - Since the problem is NP-complete [1], we utilize heuristic algorithm to solve the problem



[1] Z. Wang and J. Crowcroft, "Quality of service routing for supporting multimedia applications," IEEE Journal on Selected Areas in Communications, vol. 14, pp. 1228–1234, Sept. 1996.

Centralized heuristic algorithm

- The proposed heuristic algorithms search a set of AL routes from the available AL routes, each of which provides near-optimal performance for the AL node pair

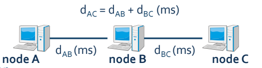
1. Centralized algorithm utilizing simulated annealing
   - The set of AL routes that are utilized by AL node pairs is regarded as state *S*
   - Neighbor state is generated by changing some AL routes of the present state to randomly generated AL routes
   - Regard a metric of expected network performance as the cost of state
     - It is calculated by the cost of AL links and routes

Algorithm 1 Centralized algorithm

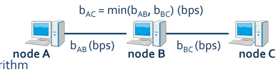
```
I ← 0,   T ← T_init,   , S ← S_init
while T > 0 do
    S_tmp ← Neighbor(S)
    if Cost(S) ≥ Cost(S_tmp) then
        S ← S_tmp
    else
        r ← Random(0, 1)
        if r < Probability(T, Cost(S), Cost(S_tmp)) then
            S ← S_tmp
        end if
    end if
    I ← I + 1
    T ← Cooling(T, I)
end while
```


How to derive the cost of AL links and routes

- Cost of AL links and routes are defined as the expected network performances when the set of AL routes is used, which calculated as below
  - End-to-end latency
    - Latency of each AL link is derived by M/M/1 queuing model using available bandwidth and link utilization



    - We regard the sum of latency of all AL links on the AL route as the end-to-end latency of the AL route
  - Available bandwidth
    - Available bandwidth of each AL route is attached by max-min algorithm



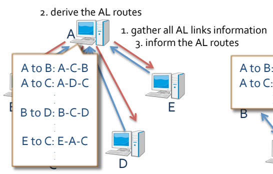
    - We regard the minimum value of all AL links on the AL route as the available bandwidth of the AL route
  - Transit cost
    - We classify IP links included by AL links as intra-AS link, peering link, and transit link
    - The transit cost of AL route is derived according to the traffic demand of the AL node pairs and kinds of inter-AS links on the AL route

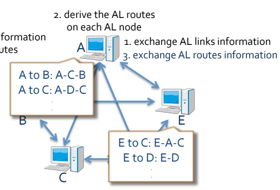
Distributed heuristic algorithm

2. Distributed algorithm utilizing distributed simulated annealing [2]
   - Run in each AL node independently
   - Each AL node measures the network performance of AL links whose source node is the AL node
   - Each AL node generates the neighbor state for simulated annealing only with the AL routes whose source node is the AL node
     - The process of distributed simulated annealing is differed from that of centralized, because the centralized algorithm targets all AL routes
   - Each AL node exchanges the information about the AL routes and network performances with other AL nodes each other
     - Each node calculates the cost of state with the information exchanged with other nodes
     - The notification of AL routes to other AL nodes makes communication overhead

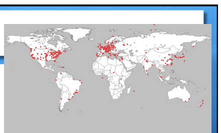
[2] M. Arshad and M. C. Silaghi, "Distributed simulated annealing and comparison to DSA," in Proceedings of the Fourth Workshop on Distributed Constraint Reasoning, Aug. 2003.

Distributed heuristic algorithm





Performance evaluation



PlanetLab web sitehttp://www.planet-lab.org/

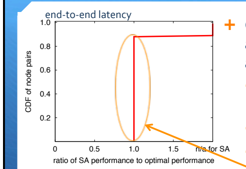
- Evaluate the proposed algorithms by assuming that the PlanetLab nodes construct an AL network and conduct an AL routing
- The process of obtaining the network performance values is as below
  - End-to-end latencies, IP-level routes
    - We conducted **traceroute** commands between all PlanetLab nodes
  - Available bandwidths, physical capacities
    - We obtained the measurement results available at Scalable Sensing Service (S3) [3]
  - AS-level routes
    - We converted the IP-level routes to AS-level routes by using the relationship information between IP address prefix and AS numbers that is available at Route Views Project [4]
  - Relationship information between ASes
    - We utilized the relationship information provided by CAIDA [5]

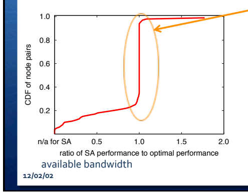
[3] "S3: A Scalable Sensing Service for Monitoring Large Networked Systems," P. Yalagandula, P. Sharma, S. Banerjee, S.-J.Lee, and S. Basu, Proceedings of the Workshop on Internet Network Measurement 2006, Pisa, Italy, September 2006.
[4] University of Oregon, "Route Views Project." available at http://www.routeviews. org/.
[5] University of California, "CAIDA." available at http://www.caida.org/home/.

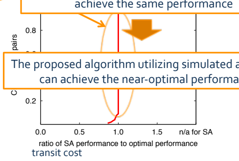
Settings for the evaluation

- Measured network performances of AL links
  - For centralized algorithm, all AL nodes obtained the network performances in advance
  - For distributed algorithm, each AL nodes obtained the network performance be assigned to itself in advance and can exchange the network performance all other AL nodes
- Value of traffic of each AL node pairs
  - 1000 kbps per an AL node pair
- Initial state of simulated annealing
  - One-hop AL routes that are equal to the IP-level routes are utilized
- Neighbor state generate function
  - Neighbor state is generated from the present state by changing 1 % of the target AL routes to randomly generated AL routes
- Exchange frequency of AL routes information
  - Exchange the AL routes information every 100 iteration of distributed simulated annealing

Comparison results with optimal performance





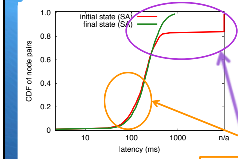


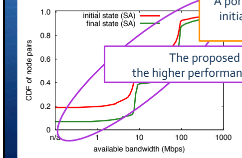
- Comparison between the optimal performance and that achieved by the proposed centralized algorithm with few AL nodes
  - 10 AL nodes are chosen from PlanetLab nodes
  - 7 AL node pairs have traffic demands
  - 100 evaluation trials are performed

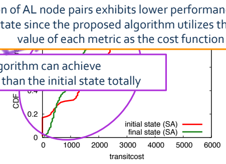
More than 70 % of all AL node pairs can achieve the same performance

The proposed algorithm utilizing simulated annealing can achieve the near-optimal performance

Performance improvement: centralized heuristic algorithm





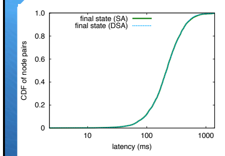


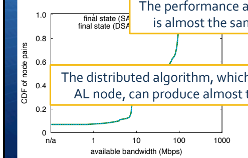
- Comparison between the initial state and the final state achieved by the proposed centralized algorithm
  - 30 AL nodes are chosen from PlanetLab nodes
  - 10 % of all node pairs have traffic demands
  - 100 evaluation trials are performed

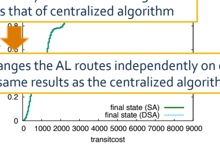
A portion of AL node pairs exhibits lower performance than the initial state since the proposed algorithm utilizes the average value of each metric as the cost function

The proposed algorithm can achieve the higher performance than the initial state totally

Performance improvement: distributed heuristic algorithm







- Comparison between the performance achieved by the centralized algorithm and that of the distributed algorithm
  - 30 AL nodes are chosen from PlanetLab nodes
  - 10 % of all node pairs have traffic demand
  - 100 evaluation trials are performed

The performance achieved by the distributed algorithm is almost the same as that of centralized algorithm

The distributed algorithm, which changes the AL routes independently on each AL node, can produce almost the same results as the centralized algorithm

Conclusion and future work

- Conclusion
  - Proposed the centralized and distributed heuristic algorithms for application-level traffic routing
  - Evaluated the proposed algorithms assuming that PlanetLab nodes performed an AL routing
    - We confirmed that the proposed algorithms could achieve near-optimal solutions as well as considerable improvement in end-to-end network performance
- Future work
  - Evaluate the proposed algorithms with more than one metrics such as minimizing end-to-end latency under a constraint on transit cost
  - Extend the proposed algorithms appropriately to the protocol developed by Application-layer Traffic Optimization (ALTO)

Ref. Formulation of application-level routing (1/2)

- Definition of IP-level routing matrix
  - Describe IP link usages by IP routes between IP routers as a matrix

Identifier of IP router pair — Identifier of IP link

$$A^{\mathrm{IP}} = \begin{pmatrix} IP_1^1 & \cdots & IP_1^{(N-1)N} \\ \vdots & \ddots & \vdots \\ IP_M^1 & \cdots & IP_M^{(N-1)N} \end{pmatrix}$$

- AL (application-level) network
  - AL nodes located at end-hosts connected to IP routers
  - AL nodes are connected each other by AL links
    - Each AL link is equal to the IP route between the AL nodes
  - AL route between AL nodes consists of one or more AL links
    - The set of available AL routes for AL node pair j, R_j^{AL} is described as follows

When AL node pair exists and it is connected by AL link, the element become one

$$(e_1^{\mathrm{AL}}, e_2^{\mathrm{AL}}, \ldots, e_{(N-1)N}^{\mathrm{AL}})$$

$$R_j^{\mathrm{AL}} = \{(p_1, p_2, \ldots, p_h) | h \geq 1,\ s_{p_1} = s_j, t_{p_h} = t_j,$$
$$t_k = s_{k+1}\ (2 \leq h,\ 1 \leq k \leq h-1),$$
$$e_{p_k}^{\mathrm{AL}} = 1\ (1 \leq k \leq h)\}$$

The equation means that AL routes that consists of available AL links

Ref. Formulation of application-level routing (2/2)

- Definition of AL routing matrix
  - The AL routing determines an AL route from available routes for each AL node pair that has traffic demand
  - Describe AL link usage as a matrix

Identifier of AL node pair

$$A^{\mathrm{AL}} = \begin{pmatrix} AL_1^1 & \cdots & AL_1^{(N-1)N} \\ \vdots & \ddots & \vdots \\ AL_{(N-1)N}^1 & \cdots & AL_{(N-1)N}^{(N-1)N} \end{pmatrix}$$

- Derive the end-to-end network performance provided by the AL routes
  - Using the AL routing matrix A^{AL} and the matrix of traffic demand between AL nodes
    - Elements of follow equations d_j^{AL} b_j^{AL} c_j^{AL} represent the end-to-end latency, available bandwidth, transit cost between AL node pair j, respectively

$$D^{\mathrm{AL}} = (d_1^{\mathrm{AL}}\ d_2^{\mathrm{AL}}\ \ldots\ d_{(N-1)N}^{\mathrm{AL}}) \quad d_j^{\mathrm{AL}} = \sum_{i=1}^{(N-1)N} d^i AL_j^i$$
$$B^{\mathrm{AL}} = (b_1^{\mathrm{AL}}\ b_2^{\mathrm{AL}}\ \ldots\ b_{(N-1)N}^{\mathrm{AL}}) \quad b_j^{\mathrm{AL}} = \min(b^i AL_j^i) \quad (1 \leq i \leq (N-1)N)$$
$$C^{\mathrm{AL}} = (c_1^{\mathrm{AL}}\ c_2^{\mathrm{AL}}\ \ldots\ c_{(N-1)N}^{\mathrm{AL}}) \quad c_j^{\mathrm{AL}} = \sum_{i=1}^{(N-1)N} c^i AL_j^i$$

performance of AL link i

Ref. Optimization problem of application-level routes

- Define the optimization problem for user-perceived performance and transit cost

  Ex) Minimize the average of the latencies of all AL routes

  Attach available route to each AL node pair

  elements of D^{AL}

$$\text{minimize}: \quad \left(\sum_{j \in K} d_j^{\mathrm{AL}}\right) / |K|$$

$$\text{subject to}: \quad r_j \in R_j^{\mathrm{AL}} \; (j \in K)$$

K : set of AL node pairs that have traffic demand

R_j^{AL} : set of AL routes available for AL node pair j

d_j^{AL} : end-to-end latency of AL route between AL node pair j

r_j : AL routes that is utilized by AL node pair j

  - If an AL route selection of each AL node pair does not affect the network performance of other AL routes, the problem is equal to that the problem to minimize/maximize the performance of each AL route independently
  - In practice, an AL route selection of each node affects the network performance of other AL routes each other
    - We adopt heuristic algorithm because the problem is NP-complete[1] and cannot be solve with exhaustive search

[1] Z. Wang and J. Crowcroft, "Quality of service routing for supporting multimedia applications," IEEE Journal on Selected Areas in Communications, vol. 14, pp. 1228–1234, Sept. 1996.